2012

Evaluating Gross vs. Net Migration Rates In a County-Level Component Model of Population Change

Henry C. Renski
University of Massachusetts - Amherst, hrenski@larp.umass.edu

Susan Strate
Donahue Institute, sstrate@donahue.umassp.edu

UMASS DONAHUE INSTITUTE

Evaluating Gross vs. Net Migration Rates In a County-Level Component Model of Population Change

Prepared for the United States Census Bureau

Henry Renski, Ph.D.,
Department of Landscape Architecture and Regional Planning
University of Massachusetts Amherst

Susan Strate,
UMass Donahue Institute

This project was funded by:
The United States Census Bureau
for Services in Support of the U.S. Census Bureau's 2010 Estimates Evaluations

Work Order: YA132310SE0381

January 2012

Contents

Tables & Figures

Overview

This paper evaluates the accuracy of county-level population estimates and forecasts under three different methods for estimating the domestic migration in a components-of-change framework. The first is a net-migration approach similar to that used by the U.S. Census Bureau and by many state data centers. While common, the net migration assumption has been widely criticized for not accurately reflecting the population 'at risk' of migrating into a county. The other two methods follow a gross migration approach whereby in- and out-migration are added separately into the population change equation. The simple gross migration approach estimates domestic in-migration to each county from the rest of the nation as a whole. The multiregional gross migration model examines flows between specific pairs of counties and adds these together to measure total in-migration. Otherwise, the population estimates models are identical – allowing us to isolate differences in population estimates due solely to how the domestic migration component is estimated.

We evaluate the accuracy of the three migration approaches against the county household population counts of the 2010 Decennial Census using a variety of common measures of predictive accuracy. We find that the simple gross migration model typically produces the smallest forecast errors. However, this is followed closely by the net migration approach, whose average forecast errors exceed the simple gross model by only .2 percentage points. Despite its far greater complexity the multiregional model produces the highest average errors of all three approaches with an average absolute error .7 percentage points higher than the net migration model. This is due largely to a higher proportion of extreme errors —counties where the model produces an average in excess of five or ten percent greater than the actual census counts. We suspect that this is due to measurement error in the Internal Revenue Service migration data, which may be more influential when calculated for specific pairs of counties but has less noticeable impact when distributed across the entire nation (i.e. the simple gross migration approach) or when in and out-migration are subtracted from one another (i.e. the net migration approach). Although producing higher errors when averaged over all counties, the multiregional model still produces the lowest errors for the greatest number of counties.

All three models produce their most reliable estimates for large counties and the greatest error for the smallest counties—places where even small differences can greatly influence year to year changes in migration rates. The simple gross migration approach is generally preferred among mid-sized and larger counties. The multiregional model is typically favored among counties with fewer than 20,000 persons. Counties experiencing rapid decline or growth are also notoriously difficult to estimate, regardless of method. Rapidly growing counties tend to be overestimated, most notably so in the case of the multiregional model which has a natural upward bias to begin with. However, the multiregional model tends to do a little better than the others at estimating population in cases of recent decline. The simple gross migration model is generally preferred for rapidly growing counties. The key exception is among fast growing small counties, which are favored by a multiregional approach.

This project was funded by:

The United States Census Bureau

For Services in Support of the U.S. Census Bureau's 2010 Estimates Evaluation

Work Order: YA132310SE0381

Background

Timely population estimates are critically important to planners and policy makers at all level of government and by many in the private and non-profit sector, as well. In the United States, the most common source for annual population estimates is the Census Bureau's Population Estimates Program which produces state and county-level estimates for the years between each decennial census. These estimates are widely used. They provide control totals for other Census Bureau estimates, such as municipal population estimates, the American Community Survey, and the Current Population Survey, to name a few. They are also commonly referenced in planning and policy documents, either when describing historical trends or serving as the foundation for forecasts. They also regularly appear as explanatory variables in scholarly studies in a variety of social and health disciplines. Many state data centers and regional planning agencies have taken to producing their own population estimates, in part to provide them with the information needed to independently verify the U.S. Census Bureau estimates or for resource distribution and planning within their own states.

A common method for measuring annual population is the components-of-change approach. Under this framework the resident population can only change through natural increase (births minus deaths) or migration. Migration can be further classified as either domestic or international. Domestic migration is the dominant source of year-to-year variation in population change for most counties in the United States (Shryock, Siegel et al. 1973; Smith 1986; Klosterman 1990). It is also typically the most volatile component of change, and as such the most difficult to estimate or forecast. Our ability to accurately estimate this component therefore hinges on our ability to getting the domestic migration component right.

There are many options for estimating domestic migration, and the choice could potentially lead to vast differences in migration estimates and forecasts. The Census Bureau, as well as most state and local analysts, currently estimate domestic migration using a net migration approach. Under the net migration approach, the migration rate is measured simply as the number of in-migrants minus the number of out-migrants, divided by the same-county population at the start of the prior year. Despite its simplicity and intuitive appeal, net migration has a serious conceptual flaw. The problem lies in the estimation of a combined (or net) migration rate for in and out-migrants (Shryock, Siegel et al. 1973; Smith 1986; Isserman 1993). Conceptually, the migration rate represents the probability of any individual moving into (or out of) a county. But for this assumption to be valid, the denominator of the migration rate must represent the population 'at risk' of migration. This is a reasonable assumption for out-migration—a county's residents in the previous year could potentially move out. But this does not hold for in-migration. The existing residents of a county are not 'at risk' of in-migration—they already live there. In fact, the population at risk of moving into a county is literally everyone but its existing residents.

While this may seem like a trivial distinction, it has the potential to create large differences in population estimates and forecasts by making growing places seemingly grow faster and declining places decline faster (Isserman 1993, 9. 47). A more theoretically valid approach is to estimate separate in- and out-migration rates and apply each to the appropriate at-risk population—existing same-county residents for out-migration and residents of the rest of the nation for in-migration. Gross migration based estimates might be improved further by modeling the flow of migrants between specific pairs of counties, rather than constructing a single in-migration rate. A multiregional model also accounts for the fact that inter-county migration is dominated by a relative small number of origins and destinations (typically large and fairly nearby counties), and presumably leads to models more sensitive to changing population dynamics of highly inter-connected areas.

The prospective benefits of implementing new population estimates must also be weighed against the costs. The argument favoring the net-migration approach is one based mainly upon practical considerations, given the realistic time, data and other resource constraints facing professional planners and demographers (Smith and Swanson 1998). The simple gross migration method requires collecting population component data for all counties in the nation, although with modern computing power this is much less of a barrier than it has been in the past. Building a true multiregional model, however, is highly data and computational intensive and requires the knowledge of analysts with a modest level of programming skill.

It is also not entirely clear that gross migration models offer a sufficient improvement in accuracy to warrant the increased time and cost. While there is a strong conceptual argument favoring gross migration approaches, there has been relatively little evaluation to document which method produces the most accurate population estimates or forecasts against actual population counts. Several studies examine the divergence of long-term forecasts produced under different assumptions against one-another. Based upon an evaluation of cohort-component forecasts for 10 states, Smith (1986) found little difference in annual estimates produced under the gross and net migration methods, overall. However, the divergence was notably higher among fast-growing and slow-growing states. Presumably, when regional population growth exceeds the nation, projections based on net rates are upwardly biased. The opposite holds in areas lagging national growth. The two methods converge when population growth in the region mirrors national trends.

Isserman (1993) is among the few to compare projections based on different migration assumptions to actual census counts. Using migration rates calculated from the 1980 Census, Isserman compared a set of 10 year cohort-component projections against the 1990 Census for the 55 counties of West Virginia. He discovered little difference between net and gross forecasts, with Mean Absolute Percentage Errors (MAPEs) of 21% and 18%, respectively. But, as noted by the author, the sample of counties may not be representative of the nation. Many West Virginian counties experienced a reversal of long-term trends in decline with the short-lived rural renaissance. This reversal could not be predicted by migration trends occurring from 1975 to 1980, as reported in the 1980 census.

In a follow-up to the 10-state 1986 study by Smith, Smith and Swanson (1998) compare cohort-based projections based on data from the 1980 Census against the actual population counts of the 1990 Census. The MAPE of the net model exceeded the gross model by 1.2 percentage points, with a 1.7 percentage point difference among the four fast growing counties in their sample. While both models tended to over-predict population in fast growing states, the upward bias of the net migration model was considerably higher than the gross migration model—with Mean Algebraic Percentage Errors (MALPEs) of 14.6% and 2.0%, respectively. Both models also underestimated the true population in low-growth states, although their relative difference is far less (-5.2% in the net migration model versus -4.1% for the gross model).

This study evaluates county household population estimates produced under three approaches to modeling domestic migration: (1) a net migration approach closely resembling the method currently used by the Census Bureau's Population Estimates Program; (2) a simple gross domestic migration approach that uses a single rest-of-the nation source as the baseline population for prospective in-migrants; and (3) a multiregional gross migration method that calculates in-migration rates between pairs of counties. We construct annual estimates for nearly every county in the U.S. from 2000 to our target year of 2010. The 100% population count of the 2010 Decennial Census provides an objective baseline to assess which of the three approaches produces the most accurate estimates according to our review of six evaluation metrics.

Calculating Net and Gross Migration Rates

Our population estimates follow a standard component-of-change framework, mirroring that used by the U.S. Census Bureau for the production of its Annual Population Estimates series.[1] The component model is based on a simple, yet conceptually appealing, demographic accounting of population change. In a given year (t) the population of any jurisdiction (i) can only change from its prior year value (t-1) through births, deaths, international migration ($IntMig$), or domestic migration ($DomMig$), or:

$$Pop_{i,t} = Pop_{i,t-1} + Births_{i,t-1} - Deaths_{i,t-1} + IntMig_{i,t-1} + DomMig_{i,t-1}.$$

In this framework, estimating population change amounts to constructing estimates of each individual component of change from available data sources.[2] This paper is mainly concerned with the estimation of the domestic migration component.

There is no direct source of information on the number of domestic migrants coming into or going out of a county. If there were, then domestic migration could be entered directly into the population equation and there would be no need for alternate estimation methods. Instead, domestic migration must be estimated from migration rates developed from proxies, which are then multiplied against the population 'at risk' of migration from the previous year. The different approaches to estimating migration (i.e. net vs. gross) come from the application of rates based on different assumptions.

---

[1] Our population estimates do not equal the official annual population estimates of the Census Bureau. One major departure is that the Census Bureau's component-based estimates are adjusted to accommodate local population challenges. The results of these challenges are not reflected in the individual components of population change and therefore could not be systematically modeled. A second difference is that the Census Bureau applies a rake factor to its county estimates so that they add-up to the exogenously estimated national population. A third difference is that the Census Bureau uses separate approaches for estimating the population above and below age 65, and then adds the two together. We were not given access to the special IRS tabulations used to create under-65 specific migration rates and therefore had to develop our model for the entire population.

[2] A common extension of the component of change model is a cohort-component model which also breaks down population change by specific sex and age (and sometimes race) cohorts. Due mainly to data limitations, this study does not expand the method to include specific cohorts of change. However, we expect our comparison of gross- and net-based population estimates to be equally valid to both component and cohort-component approaches.

Our migration rates are estimated from Internal Revenue Service (IRS) data on the total number of tax exemptions (i.e. filers and their dependents) with a change of county address from the previous year's tax return. IRS exemptions cannot be used as a direct measure of migration. Not everyone files a tax return, and some are new filers that cannot be matched to an earlier return. However, it is assumed that the rate at which exemptions move between counties are representative of actual migration patterns.

The net migration rate is simply the number of exemptions moving into county i ($InEx$) during year *t-1* minus the number exemptions moving out of county i ($OutEx$) divided by the total number of exemptions ($TotEx$). To estimate the actual number of net migrants, the net migration rate is multiplied against the base population of potential migrants (*MBase*), or:

$$NetMig_{i,t} = \left(\frac{InEx_{i,t-1} - OutEx_{i,t-1}}{TotEx_{i,t-1}}\right) MBase_{i,t-1}$$

In this case, the migration base is estimated as the number same-county residents plus half of the births, deaths, and net international migration that occurs during t-1:

$$MBase_{i,t-1} = HHPop_{t-1} + .5\left(Births_{i,t-1} - Deaths_{i,t-1} + Int\_Mig_{i,t-1}\right)$$

The gross migration approach explicitly distinguishes the forces generating domestic in-migration from out-migration, and then adds the two as separate elements into the component model of county population change. Calculating gross out-migration is very similar to calculating net migration. It is simply the share of IRS exemptions moving out of the target county i between the previous and current year multiplied by the migration base (*MBase*), or:

$$OutMig_{i,t} = \left(\frac{OutEx_{i,t-1}}{TotEx_{i,t-1}}\right) MBase_{i,t-1}$$

Calculating the gross *in-migration* rate is a bit more complicated, as the population "at risk" of moving into the target county i are persons living elsewhere in the U.S. In our simple gross migration model, we calculate gross in-migration by treating all potential counties of origin as a singular "Rest of the Nation" entity (RON), or:

$$InMig_{i,t} = \left(\frac{InEx_{i,t-1}}{TotEx_{RON,t-1}}\right) MBase_{RON,t-1}$$

where *InEx* is the number of IRS exemptions moving into the target county *i* during the past year from all other counties (*RON*). The total exemptions (*TotEx*) used in the denominator of the in-migration rate and the components of the migration base are constructed by summing over all counties, excluding *i*.

A Multiregional Model of Gross Migration Flows

Gross migration estimates can be further refined if we consider migration as a dynamic process involving the flow of people between specific pairs of specific origin and destination counties. The landscape of migration is highly uneven, with nearby and large counties typically producing more in-migrants than smaller counties that are further away. Region-specific economic conditions and similar industry profiles may also contribute to high migratory flows – for example, the Boston region may regularly trade IT workers with Silicon Valley and Austin Texas, despite their physical distance. So while our simple gross migration model may produce consistent migration rates by treating the rest of the nation as a singular entity, it may be insensitive to changes occurring in those few counties that account for the bulk of inter-regional migration.

In principle, at least, the gross migration model can be easily modified to accommodate multiple origins and destinations. The gross out-migration equation for the multiregional model is the same as it was in the simple gross-migration specification; however the number of out-migrants will differ according to changes in the *Mbase* population, which is in part based upon in-migration. The equation for in-migration is adjusted to account for flows between specific pairs of destination (i) and origin (j) counties. The in-migration from j to i ($InEx_{ji}$) is multiplied against the migration base of each j and summed over all j to estimate the number of in-migrants into i. As a practical matter, the in-migration equation must also include a RON component because the IRS does not report county-to-county migration estimates in cases where such flows are small. This residual component can be sizable, especially for small counties with few gross flows in either direction. The RON component is similar to the simple gross migration model, except it now excludes both i as well as all j counties that were covered in the county-to-county flows. The resulting equation is:

$$InMig_{i,t} = \sum_{j \neq i} \left[\left(\frac{InEx_{ji,t-1}}{TotEx_{j,t-1}} \right) MBase_{j,t-1} \right] + \left(\frac{InEx_{RONi,t-1}}{TotEx_{RON,t-1}} \right) MBase_{RON,t-1}$$

Thus, total domestic migration into target county *i* is estimated as the sum of in-migrants coming from all other *j* counties (modeled by county-to-county flows) plus a residual component reflecting the in-migrants originating in counties that are not covered by the county-to-county flows.

Data and Estimation

In order to evaluate the three different methods for calculating domestic migration, we estimate population for each year from 2001 to 2010. Like the Census Bureau, our estimates only cover the resident household population and exclude persons living in group quarters such as students living in dormitories, military personnel living in barracks, nursing homes residents, prisoners, etc.

Annual data on births, deaths, and international migration are taken from the Vintage 2009 County Components of Change files provided by the Census Bureau's Population Estimates Program. This component data covers all years from 2000 (the launch year) to 2009. As discussed previously, annual migration rates are derived from IRS tax statistics. The most recent IRS data available at the time of this study covers people moving between 2008 and 2009. We use the simple average migration rate of the three most recent years (i.e. 2006-07, 2007-08, and 2008-09) to estimate county migration rates from 2009 to 2010.

We produced annual population estimates for all counties in the United States that had consistent population and migration data over the decade. However, we had to exclude some counties from the evaluation because of major boundary changes. While most county boundaries are consistent over time, periodic changes do occur. Most often these are minor - involving a slight redrawing of boundaries that affects few (or no) households. On rare occasion, there is a major boundary change, such as when two counties are combined into one, or when a new county is carved out of one or more existing counties. Failing to account for such changes would create erroneous estimates and may lead to false conclusions.

To distinguish substantive boundary shifts from the inconsequential, we first reviewed the Census Bureau's list of all county boundary changes and corrections taking place between 2000 and 2010.[3] We then developed a GIS database of Census TIGER County boundaries in 2000 and

[3] http://www.census.gov/popest/geographic/boundary_changes/

2010, examining whether there were major boundary changes among those listed by the Census Bureau. As a rule of thumb, we focused on counties with shifts in areas of more than 1 square kilometer or where the shift affected census tracts with more than 1,000 people as of the 2000 Census. We then checked this final list against annual IRS exemptions as well as differences in the 2000 base population reported in the 2009 vintage population estimates (which accounts for boundary changes since 2000) with the 2001 vintage base (which does not). If we saw no unusual jumps in exemptions or the 2000 population estimates base, the counties were removed from the major boundary change list and included in our evaluation.

We identified 29 counties with major boundary changes. Nine are in Alaska, another nine in Nebraska, four in Texas, and the rest spread among different states. While we could not produce consistent population estimates for these 29 counties, they are still factored into the migration rates and population estimates of other counties through the RON component of the gross and multiregional migration models. Because it requires only same-county data, the inclusion or removal of these counties has no effect on the population estimates in the net migration specification.

Evaluation Metrics

The three sets of household population estimates (net, gross and multiregional) are tested against the 100% household population count of Census 2010 for a total of 3,117 counties. Our evaluation of the accuracy of each model is based upon the summary metrics requested by the Census Bureau: the Root Mean Squared Error (RMSE); Mean Absolute Percentage Error (MAPE); Mean Algebraic Percent Error (MALPE); Total Absolute Error (TAE); and Extreme Percent Error greater than plus/minus 5% and 10%.

The RMSE is the approximate average number of persons over or under-estimated by each method. The RMSE corrects for the cancelling effects of over and under-estimation through squaring, but it also tends to be more heavily influenced by prediction error in larger, more populous counties. The MAPE also accounts for offsetting tendencies of over and under-prediction, but is based on percentage error, rather than numerical error. Because it is based on percentages, the MAPE is more likely to be influenced by less populous counties that are more prone to larger percentage errors, even if the absolute number of miscounted persons is rather small. The MALPE is similar to the MAPE, but does not account for the cancelling effects of over and under-estimation. It is most commonly used to indicate the direction and relative magnitude of over or under prediction, but not the overall degree of error. The TAE is based on the absolute difference between the estimate and census count, with each measured as a share of the total (U.S.) population. Less common to evaluation studies, this measure will be higher in circumstances where large places deviate the most from the actual census count. Thus far, all of the evaluation metrics produce different variations on the notion of the average or typical error. Yet, a model that produces a slightly higher average error might actually be preferred over a model with a lower average error, if the former produces fewer outliers with very large misses. The two Extreme Percent Error measures take reliability into account by measuring the percentage of counties with an absolute percentage error greater than a specified threshold (i.e. 5% and 10%).

In addition to the measures suggested by the Census Bureau, we also include one evaluation metric of our own—the share of counties where each approach produced the most accurate estimates. For example, if the net migration model produced a lower absolute error than either the gross or multiregional approaches in 25 out of 100 counties, we say that it was most accurate for 25% of the counties. The logic of this measure is that the summary evaluation metrics (the RMSE, MAPE, etc.) generally measure the average error across all counties. But such averages are not always indicative of which model is the best for any particular county. While not accounting for the size of magnitude of errors, knowing the share of cases where each model is most accurate helps indicate whether a single model consistently produces the best results.

Results

All three estimation methods perform rather well, with generally similar estimates and degrees of estimation accuracy. Of the three methods, the simple gross migration model (Gross) is the preferred model in six of our seven evaluation metrics (Table 1). On average, the gross migration model missed the actual 2010 Census count by only 7,404 persons (RMSE) or by roughly 3% of the Census 2010 count (MAPE). The net migration method comes in a close second, with a MAPE and TAE that are both within .2 percentage points of the gross migration method. Somewhat surprisingly, the multiregional gross migration approach produced the greatest error of the three models—despite its greater complexity and more detailed modeling of county to county migration flows. On average, the multiregional model missed the 2010 Census household population count by roughly 300 persons more than the net migration approach and by 1,000 more than the gross migration model. Both the net and gross migration approaches tend to err on the side of under-estimation (MALPE), while the multiregional model is more likely to over estimate actual 2010 household population counts by a roughly similar amount.

Table 1: Net, gross, and multiregional based population estimates compared to Census 2010

	Net	**Gross**	**Mutli-Regional**
Root Mean Squared Error (RMSE)	8,168	7,404	8,493
Mean Absolute Percentage Error (MAPE)	3.2%	3.0%	3.9%
Mean Algebraic Percent Error (MALPE)	-1.3%	-1.2%	1.3%
Total Absolute Error (TAE)	2.1%	1.9%	2.8%
Extreme Percent Error (+/- 5%)	19.6%	17.3%	27.1%
Extreme Percent Error (+/- 10%)	3.9%	3.5%	6.4%
Share of Counties, Most Accurate	26.5%	33.5%	39.6%

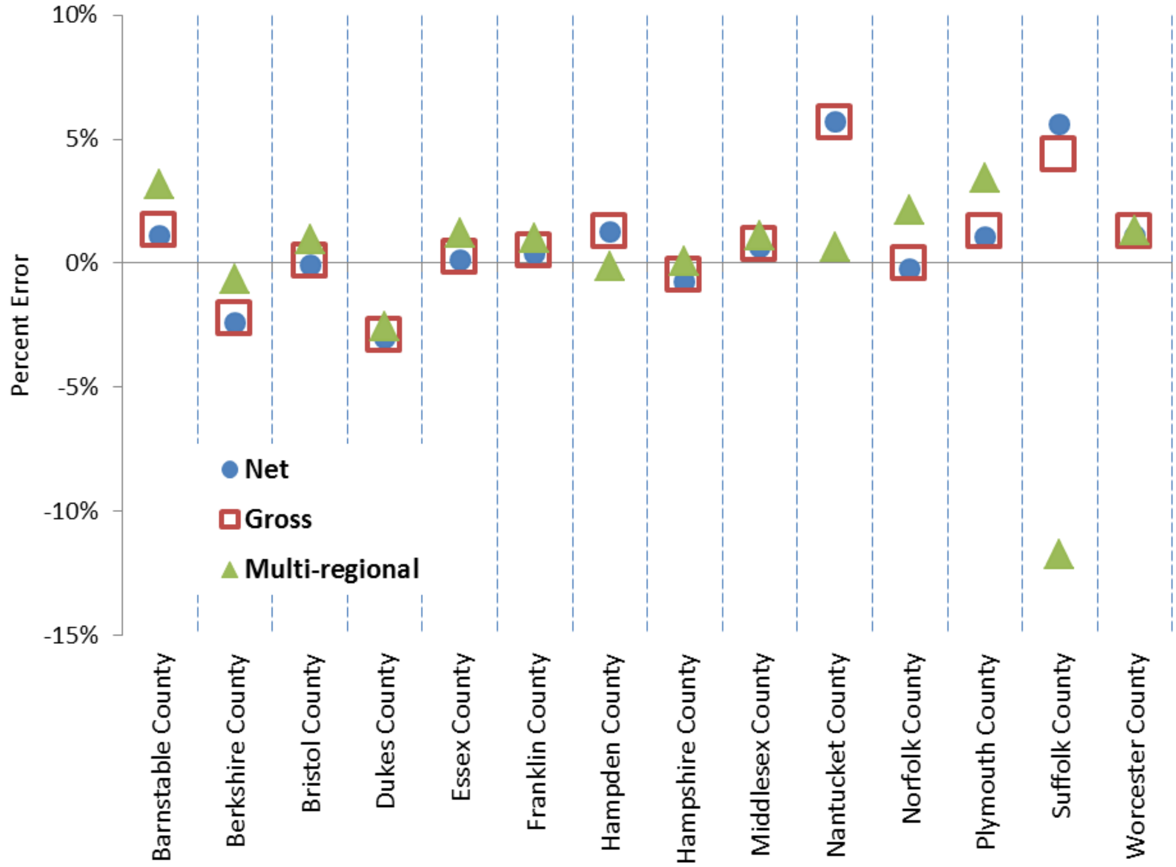
*Note: There were also 14 counties where two methods producing identical estimates, resulting in ties for the most accurate - 13 counties where gross and net estimates were tied and 1 tie between multiregional and net migration models.

The simple gross migration model also produces the fewest "extreme" errors, while the multiregional migration model produces far more. Roughly 27% of the multiregional estimates missed the Census 2010 benchmark by +/- 5%, whereas fewer than 20% of counties had errors this large when based on the net migration or the simple gross migration approach. Far fewer counties have extreme errors when the threshold is increased to greater than +/- 10% of Census 2010 counts. However, we still find that the multiregional still produces nearly twice as many counties with very high errors.

Our final evaluation metric identifies which method produced the lowest absolute error in specific counties, and then tallies that number across all counties. The results were rather unexpected. Despite its higher average error compared to the other two methods, the multiregional gross migration produced the most accurate estimates in the most counties. In just under 40% of the cases, the multiregional model produced more accurate population counts than either the net or simple gross migration methods. The gross migration model comes in a close second (34% - slightly higher if you consider ties) and the net migration model a distant third. This suggests that the higher average error of the multiregional model may not be systemic across all counties, but rather heavily influenced by places with extreme errors.

Figure 1 uses the case of Massachusetts counties to help illustrate our findings. The net and gross estimates are very close, with differences that are almost negligible in most cases. The multiregional model's estimates are also fairly close to the other two methods, although typically a little higher. The major exceptions are Suffolk and Nantucket counties, where the multiregional model produced noticeably lower estimates. These happen also to be the only counties showing extreme errors in excess of +/- 10% of the true Census count. The multiregional was far superior at estimating the Census count in Nantucket—although this is a very small county where small numerical differences may produce exaggerated percentage errors. In Suffolk County (home to Boston), the multiregional estimates were far below the Census 2010 count while the net and gross models were notably higher. It may be that the high college student population of Suffolk – which is generally not well covered by IRS migration statistics – explains the high error. Yet, other counties with large student populations - such as Hampshire - do not follow the same pattern. This suggests a need for deeper investigation into the demographic and economic conditions that are associated with the difference in estimation accuracy.

Figure 1: Percent error for net, gross and multiregional migration models, Massachusetts counties



The Influence of County Size and Population Growth on Accuracy and Bias

Population size and growth are both known to influence predictive accuracy of population estimates and forecasting methods (Smith 1987; Rayer 2008). Small areas are prone to greater error in part because their migration rates are often erratic- even small numerical changes in the number of in- or out-migrants produce notable differences in migration rates from one year to the next. Rapidly growing and declining places are also more difficult to predict. It is widely believed that the net-migration approach will exaggerate growth in fast-growing areas and population loss in declining areas (Rogers 1976; Smith 1986; Rogers 1990; Isserman 1993, 9. 47; Smith and Swanson 1998), with backing evidence offered by Smith (1986), Smith and Swanson (1998), Wilson and Bell (2004), and Isserman (1993). However, the Smith and Wilson/Bell studies are not based on comparisons to actual population counts, and the Isserman study is limited to West Virginia and does not include a multi-county, multiregional model among its comparisons.

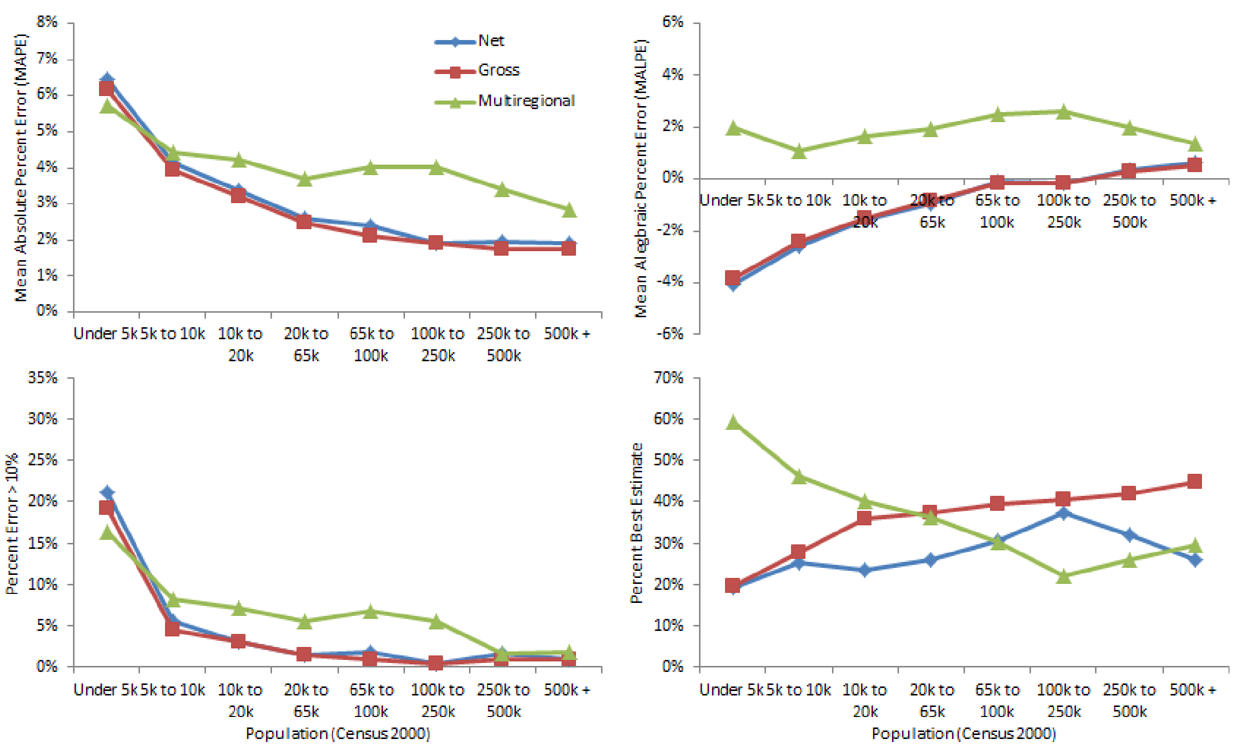
We investigate the relationship between size, growth and estimation accuracy/bias by grouping our U.S. counties by population size (as of the 2000 Census) and by growth rate, using categories provided by the Census Bureau. We then calculate our seven evaluation metrics for each, looking for consistent patterns and trends in the magnitude and direction of error. For the sake of efficiency, we focus our attention on four evaluation metrics: the MAPE, the MALPE, the percentage of counties with estimation error in excess of 10%, and the percentage of counties where each model performed best.

County Size

There is a clear relationship between size and accuracy but less evidence of consistent bias. As counties get larger, accuracy improves - as evidenced by a consistent decline in the MAPE (Figure 2). The percentage of extreme estimation errors also declines sharply with size. In fact, our analysis by size reveals the extreme errors are mainly a concern for the smallest of counties where such extreme errors represent a smaller number of people. Our measure of bias (i.e. the MALPE) shows a slight tendency of the net and gross migration methods to underestimate population more for small relative to larger counties, but this is primarily reflective of the improvement in model accuracy as county size increases.

The multiregional model stands apart from the net and simple gross migration methods in its relationship between size and accuracy. Like the others, it shows a general trend of decreasing error with size. However, the incremental improvement in accuracy is not as consistent over the size distribution and is mainly confined to the largest and smallest counties. For counties between 5,000 and 100,000 persons the multiregional model shows little substantive improvement in estimation accuracy or reduction in extreme error with size. We also see a notable tendency for the multiregional model to produce relatively more reliable estimates for the smallest counties, a tendency confirmed by our measure of the percent of counties where each model performed best. The multiregional model tends to have a fairly consistent level of positive bias, regardless of size.

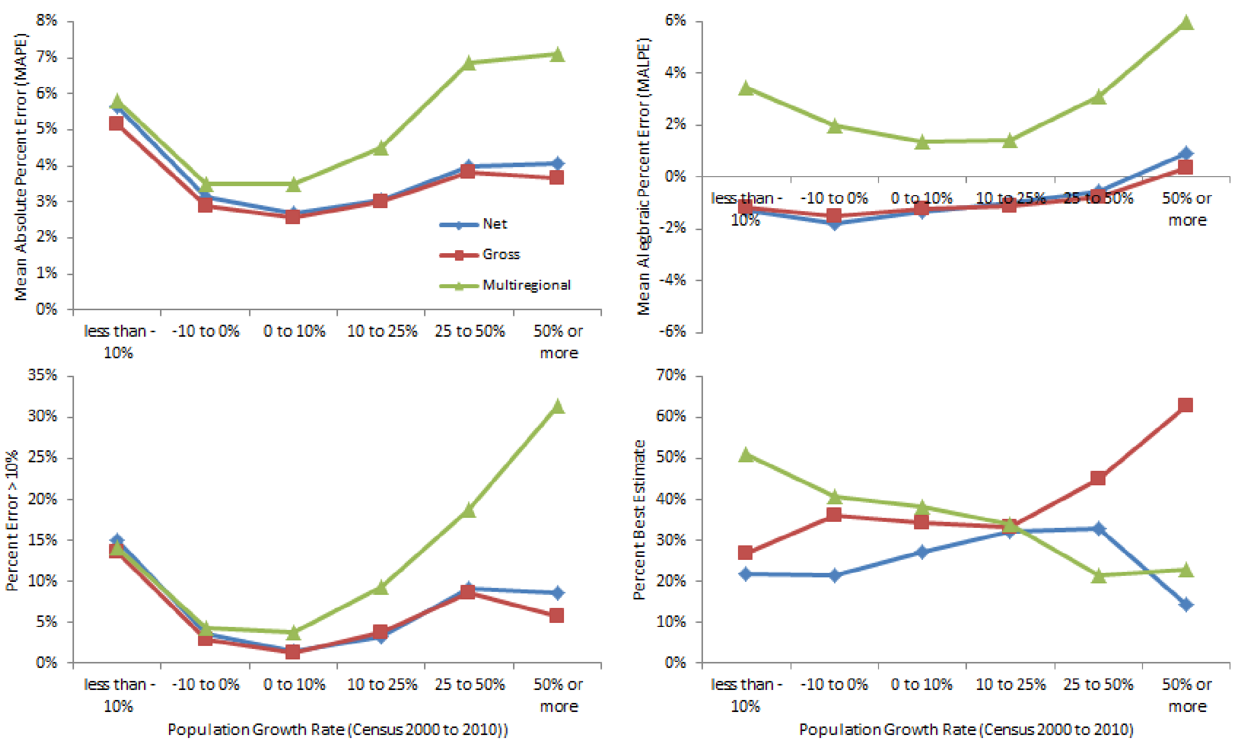
Figure 2: Summary evaluation metrics by county population size (as of Census 2000)



Population Growth

The relationship between recent population growth and accuracy is essentially non-linear. All three models have greater difficulty accurately estimating population for counties that experienced either swift decline or rapid growth over the past decade. The most reliable estimates are for counties that are stable or only slowly growing or declining. The multiregional model appears to have greater difficulty estimating the population for fast growing counties, which is also where we find the model producing extreme errors in excess of 10% of the actual census count. The negative bias of the net and gross migration models improves slightly with the rate of growth—that is the model is less likely to underestimate the true population in rapidly growing counties than it is for declining, stable or slow-growing counties. Conversely, the positive bias of the multiregional model is most pronounced for declining and fast growing counties, especially so in the case of the latter. Comparing the models directly against one another, we find that declining and slow growing regions are favored by the multivariate approach, while fast growing regions would do better adopting a simple gross migration method.

Figure 3: Summary evaluation metrics by the rate of county population growth (Census 2000 to 2010)



Population Growth Controlling for Population Size

Demographic trends aside, larger counties are less likely to register rapid growth or decline because they start from a larger base. In an attempt to distinguish the influence of population growth from county size, we repeat our analysis by examining the relationship between growth and estimation error within specific size categories. To avoid the problems associated with scarcely populated categories, it was necessary to reduce the number of population size categories from eight to four and the number of population growth classes from six to five.[4]

Table 2 summarizes the results of our examination of the relationship between estimation accuracy and bias and recent population growth, controlling for size. In general, we find that the relationship between growth and accuracy is not purely a function of differences in county size. For all but the largest size class we see the common U shaped relationship between growth and the error, whether measured by the MAPE of by the share of counties with errors exceeding 10% of the census count. All three models are better at estimating the population of stable and slow growing/declining regions, regardless of size. Error is typically the highest for the fastest growing regions except in the case of the largest counties (population of 65,000 +) where those experiencing rapid decline had substantially higher errors than even the fastest growing, although these estimates are only based upon three counties and are therefore are not reliable.

[4] The Census Bureau requested that we use the same six growth rate categories for the breakdown by growth by size as used for the analysis of growth, alone. However, there were too few counties with growth rates of 50% or higher so this category was combined with the 25% to 50% group.

Table 2: Summary evaluation statistics, growth rates within population size classes

Size/Growth	Number of Counties	MAPE Net	MAPE Gross	MAPE Multi-regional	MALPE Net	MALPE Gross	MALPE Multi-regional	Error > 10% Net	Error > 10% Gross	Error > 10% Multi-regional	Best Estimate Net	Best Estimate Gross	Best Estimate Multi-regional
Under 10,000 persons													
less than -10%	152	5.7	5.4	5.5	-2.7	-2.3	3.8	16.4	15.1	13.8	21.7	23.0	55.3
-10% to 0	279	4.6	4.3	4.3	-3.2	-2.9	1.8	9.0	6.5	8.2	21.9	28.0	47.7
0 to 10%	164	4.7	4.5	5.0	-3.1	-3.0	0.0	7.3	6.1	10.4	25.6	23.8	50.0
10 to 25%	61	5.6	5.4	5.6	-2.9	-3.2	-0.5	14.8	18.0	14.8	21.3	19.7	59.0
25% or more	16	9.6	9.8	8.9	-9.6	-9.7	-5.6	50.0	50.0	43.8	25.0	6.3	68.8
10,000 to 20,000 persons													
less than -10%	34	4.2	4.3	5.9	1.4	1.7	5.1	8.8	8.8	14.7	26.5	38.2	35.3
-10% to 0	245	3.0	2.8	3.7	-1.4	-1.2	2.5	0.8	0.8	3.7	18.4	39.2	40.8
0 to 10%	262	2.9	2.8	3.7	-1.6	-1.6	1.0	0.8	0.8	3.4	23.7	34.0	42.4
10 to 25%	90	4.1	4.1	5.4	-2.2	-2.4	0.8	7.8	7.8	16.7	34.4	32.2	32.2
25% or more	21	7.8	7.9	9.8	-5.5	-5.8	-2.1	28.6	28.6	38.1	23.8	28.6	47.6
20,000 to 65,000 persons													
less than -10%	16	5.1	4.8	5.4	2.0	1.8	2.0	6.3	6.3	6.3	18.8	31.3	50.0
-10% to 0	253	2.2	2.0	2.9	-1.1	-0.9	1.9	1.2	1.2	1.2	23.3	39.5	36.8
0 to 10%	507	2.4	2.3	3.2	-1.2	-1.1	1.7	0.6	0.8	2.8	26.6	34.9	38.1
10 to 25%	198	3.0	3.0	4.4	-0.9	-1.0	1.2	1.0	1.0	10.6	26.8	34.8	38.4
25% or more	72	3.8	3.5	7.7	1.1	0.8	5.6	6.9	5.6	25.0	30.6	55.6	12.5
65,000 persons or higher													
less than -10%	3	20.6	8.1	22.7	20.6	8.1	-22.0	66.7	33.3	66.7	0.0	66.7	33.3
-10% to 0	98	1.7	1.6	2.5	-0.2	-0.1	1.4	2.0	2.0	3.1	23.5	43.9	32.7
0 to 10%	300	1.9	1.8	3.1	-0.3	-0.2	1.9	0.7	0.0	2.7	31.7	39.3	29.0
10 to 25%	234	2.0	1.9	4.0	-0.1	-0.2	2.4	0.4	0.9	3.8	38.9	36.3	24.8
25% or more	112	2.6	2.4	5.6	1.1	0.8	4.7	0.9	0.0	11.6	31.3	52.7	16.1

We again see the familiar result of multiregional model producing more reliable estimates for small counties, whereas the net and simple gross migration models work better for mid-sized and larger counties. When stratified by specific size classes, we find the multiregional model is particularly adept at estimating the population of fast growing small counties and far worse at accurately estimating population of fast growing large counties. The gross and net migration models produce similar estimates with the gross migration model generally preferred. According to the MALPE, all three models increasingly underestimate population in smaller counties that are growing rapidly. This may help explain why the multiregional model, which is more prone to systematic overestimation, does better among this group.

Summary and Discussion

In this study we test which of three methods for estimating domestic migration (net, gross and multiregional gross) is best at predicting the actual household population counts reported in the decennial Census of 2010. Overall, the simple gross migration-based method produces lower average prediction errors than the net migration method currently used by the Census Bureau and many state and regional data centers and planning agencies. The simple gross migration model also produces the fewest counties with "extreme" prediction errors in excess or 5% and 10%. However, the relative improvement in predictive power is fairly small, typically under 1,000 persons (on average) or within a single percentage point of the net migration-based population estimates.

The bulk of our research effort was developing a true multiregional model of gross migration flows, where county in-migration is based on migration rates calculated between county pairs, rather than as a national aggregate. This approach acknowledges that much domestic migration is intra-regional, with a few (often adjacent) counties accounting for the bulk of migratory flows. A multiregional framework enables the analyst to look more closely at the dynamics of these highly interrelated counties, thus leading to more informed and (presumably better) models.

Despite its use of more detailed migration data, our multiregional model produced population estimates with the highest average prediction errors of the three approaches tested. While the net and simple gross migration models tend to underestimate true population counts, the multiregional model tends to err on the side of over-estimation. The multiregional model also had a higher percentage of counties with extreme estimation errors. Our results do not definitely rule out the value of a multiregional approach, however. In fact, the multiregional model actually produced more accurate estimates than either the simple gross and net-migration models for most counties. But when the multiregional model misses the mark, it tends to miss by a lot – raising its overall average prediction error and resulting in a greater share of counties with extreme errors.

At this juncture, we can only speculate on why the multiregional model produced more extreme errors in these counties. One likely candidate is measurement error in the IRS exemptions data used to estimate migration rates. The multiregional model is likely to be much more sensitive to measurement error in the IRS migration data than the other two approaches because it is based on flows between fairly small spatial units. In the simple gross migration model, county-specific measure error gets aggregated into a rest of nation residual, and thus also has little effect on the aggregate gross in-migration rate. In the net migration method, errors in estimates of in-coming migrants may be somewhat offset by subtracting out-going migrants. In future work, we hope to explore this issue further.

Consistent with previous research, we find that all three models are best at estimating populations for larger counties that are relatively stable. Fast growing and rapidly declining counties are notoriously difficult to estimate, as are very small counties. Comparing the three models against one another, we find the multiregional model to be the model of choice among very small counties and in cases of recent population decline. Likewise, the simple gross migration model is preferred in cases of mid-sized and large counties, as well as for rapidly growing counties. The key exception is among small counties experiencing rapid growth – where the multiregional model provides more reliable estimates.

In conclusion, we recommend that the Census Bureau's Population Estimates Program consider developing and testing a county population estimates model based on a simple gross migration approach. Considering that the population estimates for gross and net migration are very close, it is possible that a formal evaluation by the Census Bureau could come to a different conclusion. While we tried our best to imitate the current Census approach, data limitations ultimately led us to make a few minor adjustments. Of greatest importance is that our population estimates cover the total household population as a whole. By contrast, the Census Bureau develops separate estimates for persons above and below 65 years of age, in which case the IRS migration data is only used as a source for modeling migration of the under-65 group.[5] It is possible that the net migration specification might be actually favored if only applied to the under-65 population. But it is also possible that the gross migration model might prove to be even more effective if the 65 years and older group were removed. Only through explicit testing can we know for sure.

The benefits of implementing a more accurate model must ultimately be weighed against the costs—namely the time and effort spent developing, implementing and testing a new approach. On the one hand, the gross and net based migration estimates produce generally similar estimates. On the other hand, the costs of implementing this approach are small. The simple version of the gross migration model is not that much more complicated than the net model, with the only major difference being the need to tabulate the total number of in-migrants, total exemptions, and baseline migrant population for the nation as a whole and then subtracting activity in the destination county. In the past, tabulating IRS exemption data for the nation may have been a practical barrier for states and regional planning agencies developing their own estimates—but not anymore. Any desktop computer with a common spreadsheet application could easily handle this task. Furthermore, the IRS now makes its county migration data freely available on its website.

Given its higher average error combined with its considerable complexity, we do not recommend adoption of the multiregional model as the single approach favored by the Census Bureau. However, this approach might still be favored by regions and states conducting their own independent population estimates and forecasts. No estimation method produces superior estimates in all circumstances. Rather than simply picking the method that produces the lowest average error across all types of counties, it is perhaps more valuable to know which estimation method is more appropriate to the circumstances faced by a particular region. In this study we developed a multiregional population estimates model for nearly every county in the nation. In doing so, we were unable to scrutinize and refine the results for individual counties beyond an examination of major outliers and our checks for the internal consistency of the model. The

[5] We originally requested the special (under 65 years of age, only) tabulations of the county-to-county IRS migration data for the entire nation. Because we were unable to obtain these data, our migration estimates instead include all age groups. Because we were unable to obtain these data, our migration estimates instead include all age groups.

multiregional model would be much more tractable if created for a smaller number of counties where the analyst could pay greater attention to the influence of individual counties on domestic migration. The multiregional model would be particularly valuable in a forecasting context, where it could be used to develop a range of likely population forecasts based on alternate scenarios for growth and decline of the most closely connected counties.

Works Cited

Isserman, A. M. (1993). "The Right People, The Right Rates - Making Population Estimates with an Interregional Cohort-Component Model." Journal of the American Planning Association **59**(1): 45-64.

Klosterman, R. (1990). Community Analysis and Planning Techniques. Savage, MD, Rowman & Littlefield.

Rayer, S. (2008). "Population Forecast Errors." Journal of Planning Education and Research **27**(4): 417-430.

Rogers, A. (1976). "Shrinking large-scale population-projections models by aggregation and decomposition." Journal of Regional Science **9**: 417-424.

Rogers, A. (1990). "Requiem for the Net Migrant." Geographical Analysis **22**(4): 283-300.

Shryock, H., J. Siegel, et al. (1973). The Methods and Materials of Demography. Washington DC, U.S. Bureau of the Census.

Smith, S. K. (1986). "Accounting for Migration in Cohort-Component Projections of State and Local Populations." Demography **23**(1): 127-135.

Smith, S. K. (1987). "Tests of forecast accuracy and bias for county population projections." Journal of the American Statistical Association **82**: 991-1003.

Smith, S. K. and D. A. Swanson (1998). "In defense of the net migrant." Journal of Economic & Social Measurement **24**(3/4): 249-264.

Wilson, T. and M. Bell (2004). "Comparative empirical evaluations of internal migration models in subnational population projections." Journal of Population Research **21**(2): 127-160.